Predictions of 2010's tropical cyclones using the GFS and ensemble-based data assimilation methods

Thomas M. Hamill[1], Jeffrey S. Whitaker[1], Daryl T. Kleist,[2]

Michael Fiorino[3], and Stanley G. Benjamin[3]

[1] *NOAA Earth System Research Laboratory, Physical Sciences Division, Boulder, CO*

[2] *Environmental Modeling Center, National Centers for Environmental Prediction, Camp Springs, MD*

[3] *NOAA Earth System Research Laboratory, Global Systems Division, Boulder, CO*

Submitted as an "expedited contribution" to
Monthly Weather Review

4 May 2011

Corresponding author:

Dr. Thomas M. Hamill
NOAA ESRL, Physical Sciences Division
R/PSD 1, 325 Broadway
Boulder, CO 80305-3328
Tom.hamill@noaa.gov
Phone: (303) 497-3060
fax: (303) 497-6449

ABSTRACT

Experimental ensemble predictions of tropical cyclone (TC) tracks from the ensemble Kalman filter (EnKF) using the Global Forecast System (GFS) model were recently validated for the 2009 Northern Hemisphere hurricane season (Hamill et al. 2011). A similar suite of tests is described here for the 2010 season. Two major changes were made this season: (1) a reduction in the resolution of the GFS model, from 2009's T384L64 (~31 km at 25°N) to 2010's T254L64 (~47 km at 25°N), and some changes in model physics, and (2) the addition of a limited test of deterministic forecasts initialized from a hybrid 3-D Variational/EnKF method.

The GFS/EnKF ensembles continued to produce reduced track errors relative to operational ensemble forecasts created by the National Centers for Environmental Prediction (NCEP), the United Kingdom Met Office (UKMO), and the Canadian Meteorological Centre (CMC). The GFS/EnKF was not uniformly as skillful as the European Centre for Medium Range Weather Forecast (ECMWF) ensemble prediction system. GFS/EnKF track forecasts had slightly higher error than ECMWF at longer leads, especially in the western North Pacific, and exhibited poorer calibration between spread and error than in 2009, perhaps due in part to lower model resolution. Deterministic forecasts from the hybrid were competitive with deterministic EnKF ensemble-mean forecasts and superior in track error to those initialized from the operational variational algorithm, the Grid-Point Statistical Interpolation (GSI). Pending further successful testing, NOAA intends to implement the global hybrid system operationally.

1. Introduction.

Recently, Hamill et al. (2011; H11 hereafter) verified global ensemble predictions of 2009's Northern Hemisphere summer tropical cyclone forecasts initialized with a global ensemble Kalman filter (EnKF) system. The assimilation and forecasts were performed at a relatively high resolution, T384L64, or approximately 31 km at 25°N[1]. The significant improvement of these experimental forecasts relative to the operational global ensemble guidance provided by NCEP and other centers motivated us to continue experimentation with a similar system during the 2010 Northern Hemisphere hurricane season. More constrained computational resources in 2010 led us to choose to test a somewhat reduced-resolution system (T254L64, ~ 47 km at 25°N). However, a more recent version of the GFS model was also used[2], and a larger sample of forecasts was produced in

---

[1] The computation of the equivalent grid spacing for a given triangular truncation can be somewhat ambiguous. A nominal equivalent grid spacing depends on whether the model forecast is transformed from its spectral basis to a linear or to a Gaussian grid. A truncation at wavenumber M on a linear grid (ECMWF's approach) uses 2M+1 grid points around a latitude circle; a Gaussian grid (NCEP's approach) uses 3M+1 grid points. Use of the Gaussian grid eliminates aliasing errors during the computation of the products of two or more truncated spectral harmonic expansions but otherwise does not provide a finer discretization and more resolution. For more information, see Durran (1998, section 4.4.3). The equivalent grid spacing calculations here assume 111 km/degree * cos(latitude) for the number of kilometers per degree at a given latitude.

[2] The web page http://www.nco.ncep.noaa.gov/pmb/changes/ provides a summary of the changes to the GFS model. The 2009 season EnKF forecasts were conducted with the model version with an implementation date of Dec 15, 2009, and the 2010 season EnKF and hybrid forecasts were conducted with the model version with the implementation date of July 28, 2010. More detail on the model changes for the 2010 season are provided at http://www.nws.noaa.gov/os/notification/tin10-15aab_gfs.txt .

2010. 0000 UTC and 1200 UTC 5-day forecasts were generated each day from early June 2010 through to the end of October.

A "hybrid" 3D-Variational (3D-Var) and EnKF system was also tested during 2010. The T254 EnKF was cycled alongside the T574 Global Statistical Interpolation System (GSI; Kleist et al. 2009a). The static background-error covariances used by the GSI were blended with EnKF covariances for the GSI assimilation, 2/3 ensemble and 1/3 static. The blending used the α–control variable approach described in Lorenc (2003) and Wang et al. (2007). Before making short-term ensemble forecasts for the next update, the EnKF perturbations were re-centered on the hybrid analysis. For the EnKF, a horizontal localization length scale (i.e., the length at which covariances taper to zero) of 1600 km was used. The vertical length scale was 1.6 units in the natural logarithm of pressure. Similar length scales were used in the variational analysis for the localization of the control variable used to apply the ensemble contribution to the background error. As described in H11, advisory minimum sea-level pressure observations ("TCVitals") were assimilated by both the EnKF and the hybrid. The tangent-linear normal mode constraint (Kleist et al. 2009b) was applied to the full analysis increment (combined ensemble and static components) in the hybrid analysis.

Below, we briefly review the performance of the EnKF and hybrid systems for track error. The reader is referred to H11 for background on the forecast models, the data assimilation systems, and the verification techniques used here. Note two significant changes to the operational ensemble systems since H11. First, the ECMWF increased the horizontal resolution of their ensemble prediction system

(EPS) to T639L62 resolution (~28 km at 25°N; Miller et al. 2010), and ECMWF in 2010 added initial-condition perturbations generated from a perturbed-observation 4D-Var to their normal singular-vector perturbations (Isaksen et al. 2011, Buizza et al. 2011). Second, NCEP increased the resolution of its EPS to T190L28 (~63 km at 25°N; http://www.emc.ncep.noaa.gov/gmb/ens/ens_imp_news.html). The operational NCEP ensemble used version 8.0 of the GFS, implemented in July 2007. The CMC ensemble system continued to use a horizontal computational grid of 400x200, or approximately 0.9 degrees, and 28 vertical levels. The UKMO ensemble system was run with a grid spacing of 1.25 degrees longitude and 0.83 degrees latitude, with 38 vertical levels.

2. **Performance of the GFS/EnKF relative to operational ensembles.**

Figure 1 provides homogeneous comparisons for 2010 of average track error of the ensemble-mean track and the spread in the ensembles, following the definitions for these from H11. The GFS/EnKF provided statistically significant lower track errors than the operational GFS ensemble and exhibited a greater consistency between spread and error, though the GFS/EnKF was less consistent than it was for the higher-resolution forecasts tested for the 2009 season (H11, Fig. 5), perhaps due to the reduced resolution (2010's T254 vs. 2009's T382). Compared with ECMWF, the GFS/EnKF forecasts were slightly higher in error at the longer leads, and the spread magnitude was not as consistent with the error magnitude. While the 2009 GFS/EnKF and ECMWF systems were comparable in horizontal resolution, T384 and T399 respectively, in 2010 the resolutions were

T254 and T639 respectively, a dramatic advantage for ECMWF. GFS/EnKF forecasts were also lower in error than CMC and UKMO forecasts.

Figure 2 shows that the relative performance differed between basins. ECMWF had a relatively consistent error across all basins, while the GFS/EnKF had smaller errors than ECMWF in the Atlantic and Eastern Pacific basins and the larger errors ECMWF in the Pacific basin; these qualitative differences also occurred in 2009 (not shown). The degraded performance in the western Pacific was not the result of a few badly busted GFS/EnKF forecasts, as may occur when a TC recurvature was mis-forecast; 41 of the 58 three-day forecasts had lower error in the ECMWF system. Further, the underestimate of forecast uncertainty in the GFS/EnKF ensemble primarily happened for those forecasts with large spreads, indicating that an even larger spread in the track forecasts was warranted (not shown).

The reason for the variable performance of the GFS/EnKF in different basins and the higher error in the western Pacific was not clear. The GFS/EnKF did utilize TCVitals estimates of cyclone location and central pressure that ECMWF did not. It is possible that this data was of higher quality in the Atlantic basin, where more regular reconnaissance flights were conducted. Another possibility is that steering winds in the EnKF were more poorly forecast in the western Pacific. Figure 3 shows differences in 200-hPa mean wind analyses between the EnKF and ECMWF. Differences were somewhat more pronounced in the western Pacific than they were in the Atlantic, and were largest in the monsoon region of the Indian Ocean. Perhaps the ECMWF analyses were notably better in the western Pacific and Indian Oceans due to their recently improved forecasts of monsoon dynamics (Chakraborty 2010) and/or the Madden-Julian Oscillation (Vitart and Molteni 2011).

3. Comparative performance of the hybrid deterministic forecasts.

To evaluate the hybrid GSI/EnKF system, three data assimilation systems were compared. The first system was the operational "GSI" system. The data assimilation and deterministic forecasts were made with the GFS at T574L64 resolution (~21 km at 25°N). For the second, the "EnKF," T574L64 deterministic GFS forecasts were generated by initializing from the T254L64 EnKF ensemble-mean analysis. The third was the hybrid system, which performed the data assimilation and deterministic forecast at T574, though the input ensemble information was at T254L64. For this comparison, computations were performed only for one active period during the season, from 26 July to 7 September 2010.

Figure 4 shows homogeneous comparisons of the hybrid data assimilation system with respect to the GSI and the EnKF. The hybrid had track errors that were a statistically significantly improved from those from the GSI but were nearly indistinguishable from those provided by the EnKF. Interestingly, Fig. 5 shows that 72-h deep-layer tropical wind forecasts were lower in error (relative to the ECMWF analysis) in the EnKF than in the GSI, and the hybrid errors were even lower. Why the reduced hybrid wind errors did not result in improved track forecasts is still being explored. The relative forecast performance from the various initializations was more similar in the mid-latitudes (not shown), which may explain some of the lack of difference.

4. Conclusions.

In 2010, the EnKF system continued to produce improved track forecasts relative to the NCEP operational system (which used an older model with different initialization and reduced resolution). The hybrid data assimilation system, now in development, showed comparable track errors to the EnKF. During the next year, the hybrid system will be further refined. Pending a satisfactory evaluation of the performance of forecasts using a variety of metrics, the hybrid method is expected to become the operational data assimilation system for NCEP.

Acknowledgments.

This research has been funded by (1) NOAA's THORPEX program; (2) the NOAA Hurricane Forecast Improvement Program. We are grateful for the work of Dave Parrish (NCEP) for leading the implementation of the hybrid software in GSI, John Derber (NCEP) for optimization of the EnKF, and Xuguang Wang (University of Oklahoma) for contributions to the hybrid development.

References

Buizza, R. Leutbecher, M., Isaksen, L., J. Haseler, 2011: Combined use of EDA and SV-based perturbations in the EPS. *ECMWF Newsletter 123*, 22-28.

Chakraborty, A., 2010: The skill of ECMWF medium-range forecasts during the year of tropical convection 2008. *Mon. Wea. Rev.*, **138**, 3787–3805. doi: 10.1175/2010MWR3217.1 .

Durran, D., 1998: *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics*. Springer Press, 482 pp.

Hamill, T. M., J. S. Whitaker, M. Fiorino, and S. J. Benjamin, 2011. Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev*., **139**, 668-688.

Isaksen, L., J. Haseler, R. Buizza, and M. Leutbecher, 2011: The new ensemble of data assimilations. *ECMWF Newsletter 123*, 17-21.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. J. Lord, 2009a: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691-1705.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009b: Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.*, **137,** 1046-1060.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.,* **129,** 3183–3203.

Miller, M., R. Buizza, J. Haseler, M. Hortal, P. Janssen, and A. Untch, 2010: Increased resolution in the ECMWF deterministic and ensemble prediction systems. *ECMWF Newsletter 124*, 10-16.

Vitart, F., and F. Molteni, 2011: Simulation of the Madden-Julian Oscillation and its impact over Europe in ECMWF's monthly forecasts. *ECMWF Newsletter 126*, 12-17.

Wang, X, C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently proposed ensemble–3DVAR hybrid analysis schemes. *Mon. Wea. Rev.,* **135**, 222-227.

FIGURE CAPTIONS

Figure 1: Homogeneous comparison of global average absolute track forecast errors and average spread between the experimental GFS/EnKF and (a) the NCEP operational ensemble system, (b) the ECMWF operational ensemble system, (c) the CMC operational ensemble system, and (d) the UK Met Office operational ensemble system. Numbers in parentheses up top on each panel indicate the sample size at a particular forecast lead, i.e., the number of matched paired forecasts between the GFS and the model in question. Dashed lines indicate spread, solid lines indicate error. Error bars indicate the 5th and 95th percentiles of a resampled block bootstrap distribution as described in section 2d of H11.

Figure 2: As in Fig. 1, but for the GFS/EnKF vs. the ECMWF model in the (a) western Pacific basin, (b) eastern Pacific, and (c) Atlantic.

Figure 3: Differences between EnKF and ECMWF 200 hPa wind analyses, averaged from the period of 0000 UTC 26 July 2010 to 1200 UTC 27 September 2010.

Figure 4: Homogeneous comparison of deterministic track errors from T574 hybrid relative to (a) GSI operational, and (b) experimental EnKF.

Figure 5: Deep-layer (850 to 200-hPa average) mean vector wind errors of the GSI 3D-Var, the EnKF, and the hybrid systems for the period 0000 UTC 26 July 2010 to 1200 UTC 17 September 2010. Black curves denote the GSI, red curves the EnKF, and blue curves the hybrid. Tropics here refer to the latitude band from 20° N to 20°S.

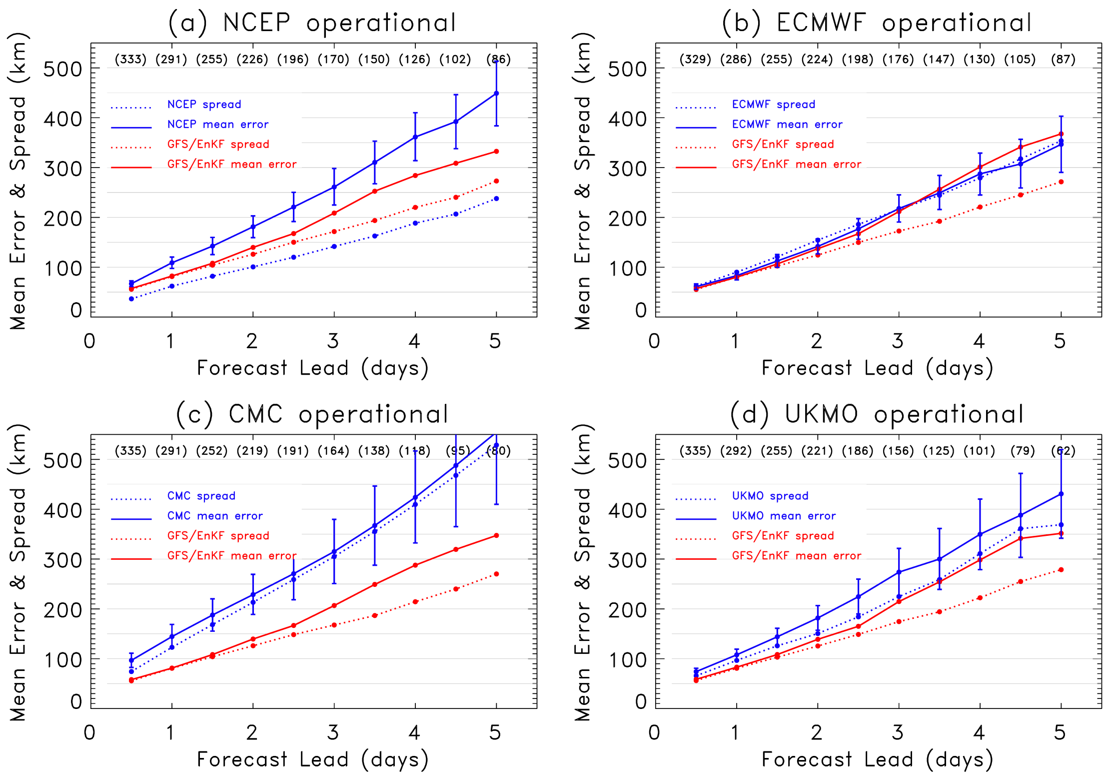

Figure 1: Homogeneous comparison of global average absolute track forecast errors and average spread between the experimental GFS/EnKF and (a) the NCEP operational ensemble system, (b) the ECMWF operational ensemble system, (c) the CMC operational ensemble system, and (d) the UK Met Office operational ensemble system. Numbers in parentheses up top on each panel indicate the sample size at a particular forecast lead, i.e., the number of matched paired forecasts between the GFS and the model in question. Dashed lines indicate spread, solid lines indicate error. Error bars indicate the 5th and 95th percentiles of a resampled block bootstrap distribution as described in section 2d of H11.

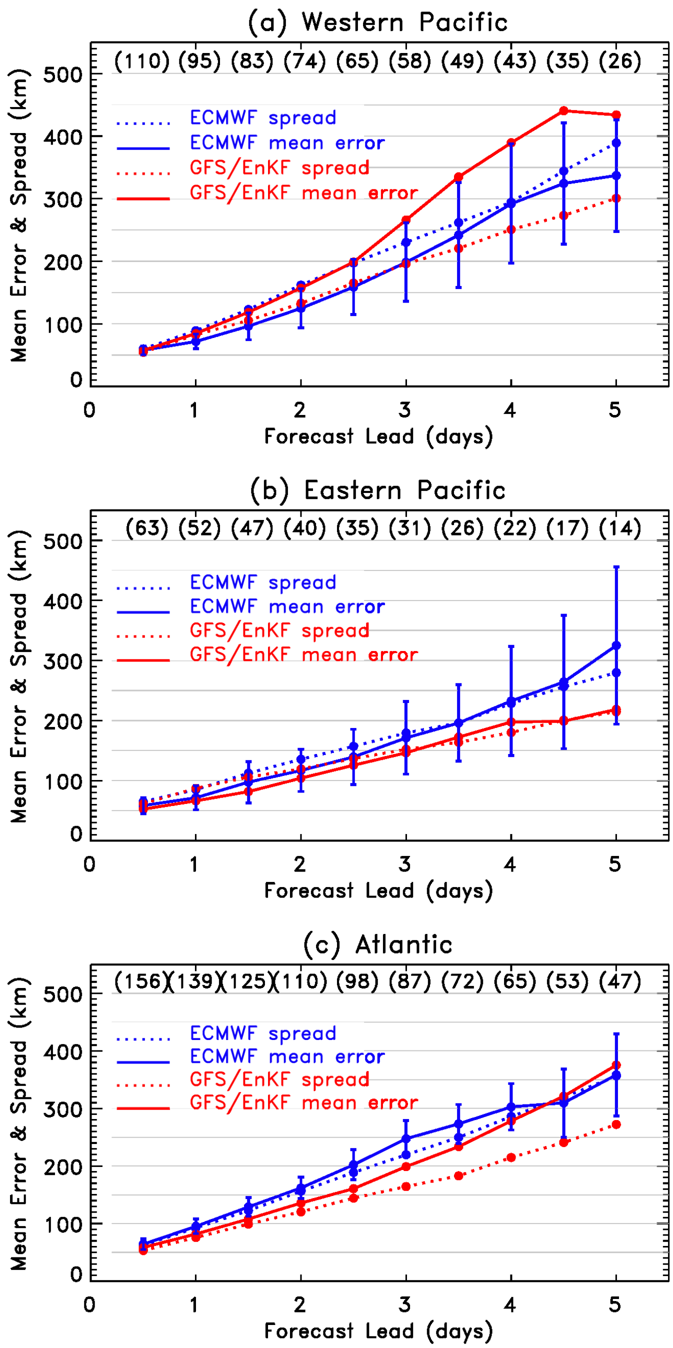

Figure 2: As in Fig. 1, but for the GFS/EnKF vs. the ECMWF model in the (a) western Pacific basin, (b) eastern Pacific, and (c) Atlantic.

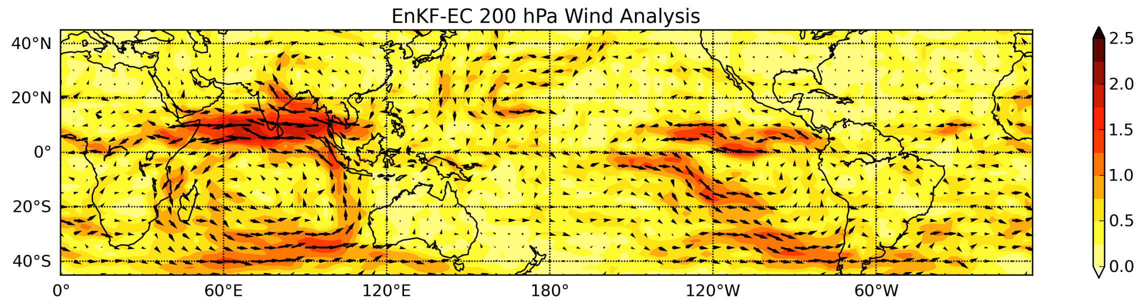

Figure 3: Differences between EnKF and ECMWF 200 hPa wind analyses, averaged from the period of 0000 UTC 26 July 2010 to 1200 UTC 27 September 2010.

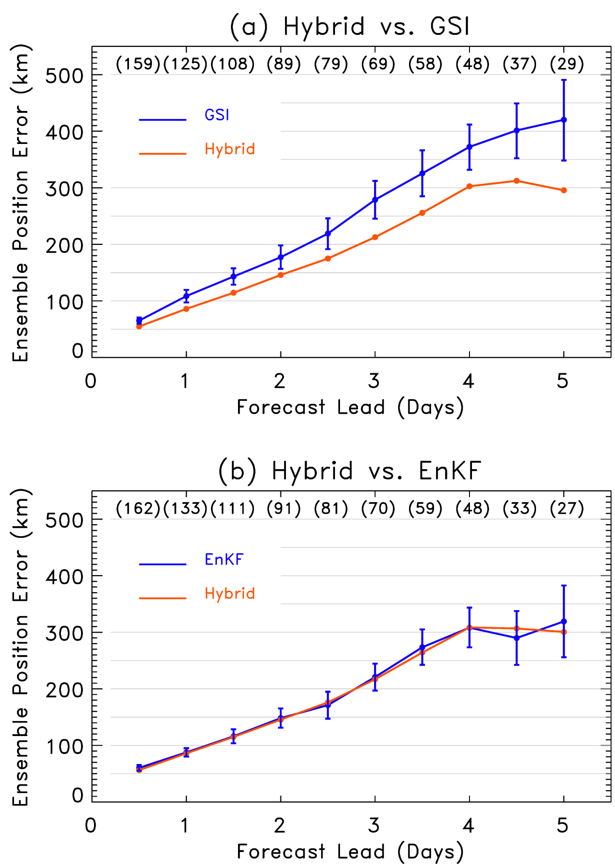

Figure 4: Homogeneous comparison of deterministic track errors from T574 hybrid relative to (a) GSI operational, and (b) experimental EnKF.

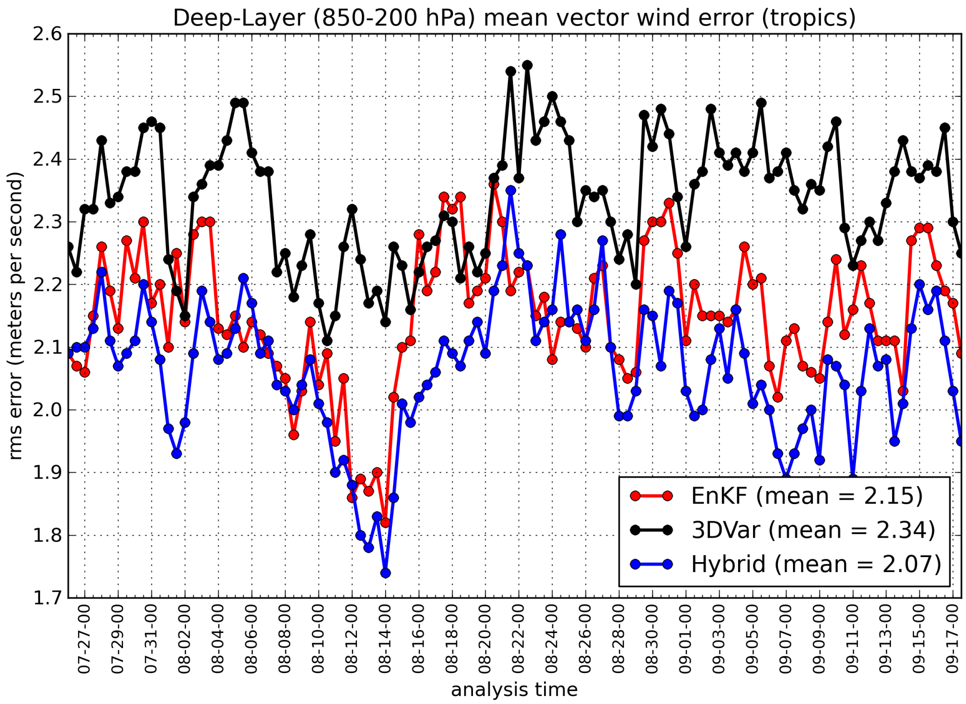

Figure 5: Deep-layer (850 to 200-hPa average) mean vector wind errors of the GSI 3D-Var, the EnKF, and the hybrid systems for the period 0000 UTC 26 July 2010 to 1200 UTC 17 September 2010. Black curves denote the GSI, red curves the EnKF, and blue curves the hybrid. Tropics here refer to the latitude band from 20° N to 20°S.